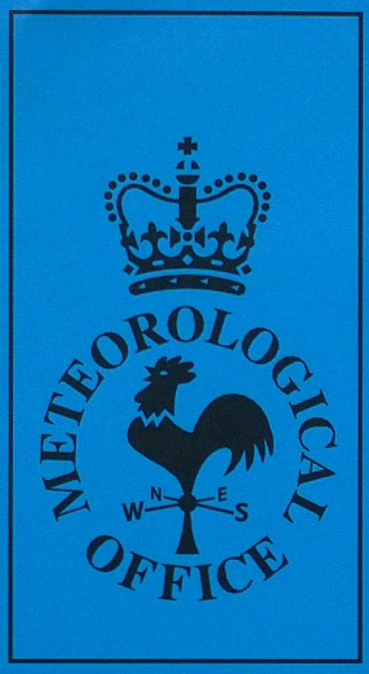

Forecasting Research

Forecasting Research Division
Technical Report No. 197

USE OF HIGH RESOLUTION LAND USE DATA IN THE UK MESOSCALE MODEL

by

C P Jones

August 1996

Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

Forecasting Research
Technical Report No. 197

USE OF HIGH RESOLUTION LAND USE DATA IN THE UK MESOSCALE MODEL

by

C P Jones

August 1996

Forecasting Research
Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

© Crown Copyright 1996

This paper has not been published and should be regarded as an Internal Report from the Meteorological Office. Permission to quote from it should be obtained from the above Meteorological Office division.

ACKNOWLEDGMENTS

Copyright and ownership of the NATLAC database is held by the Natural Environment Research Council. Further details can be obtained from: Environmental Information Centre, ITE Monks Wood, Abbots Ripton, Huntingdon Cambs.

I would also like to thank Byron Chalcraft for helping to choose suitable model case studies and Steve Woltering for assistance in using the Mesoscale Model station verification system.

ABSTRACT

The surface physics parametrizations require accurate knowledge of the underlying surface type and vegetation cover in order to calculate fluxes of heat, moisture and momentum between the atmosphere and the underlying surface. These fluxes then influence the near surface parameters of temperature, humidity and wind speed.

Until recently, the only dataset available has been one produced for climate models and does not provide the spatial detail required for operational numerical weather prediction models. This paper describes the development and testing of a much higher resolution dataset for Great Britain. The processing of the data is described and a series of case studies run to examine the impact of the new data.

CONTENTS

1. Introduction....4
2. The WHS dataset....5
3. ITE NATLAC dataset....6
4. Processing the ITE dataset....8
5. Using the ITE dataset....19
6. Case studies....30
7. Conclusions....41
8. References....41

1. INTRODUCTION

The importance of the underlying surface type in atmospheric general circulation models is increasingly being recognized. Fluxes of heat, moisture and momentum between the atmosphere and the underlying surface are a function of the physical properties of the surface type. These fluxes are important in determining the near-surface parameters of temperature, humidity and momentum. Imagine being on open moorland compared to being in a dense forest in a strong gale.

As the horizontal resolutions of models increases and parametrizations of sub-grid scale heterogeneity are developed it will become more and more important to have accurate knowledge of the surface type.

The current surface type dataset used in GCMs worldwide, including the Unified Model (UM) at the UK Meteorological Office, is that created by Wilson and Henderson-Sellers (1985), hereafter referred to WHS. This dataset was created originally for climate models and the data is presented on a regular 1°x1° latitude-longitude grid. This resolution is adequate for climate modelling but is insufficient for operational numerical weather prediction especially for Mesoscale Modelling. Unfortunately, in the absence of any other data, the UM has been forced to use this dataset.

Recently, a high resolution dataset of surface land use has been obtained for an area comprising of England, Scotland and Wales. This dataset has been used to create a new dataset for the UK Mesoscale Model. This paper details the processing of the data and results of case studies showing the impact of the data.

It is also worth noting that the UM surface parametrizations have been developed to use the WHS dataset. This work has demonstrated that it is possible to import other data into the system although some manipulation may be required. This work is an important prototype to hopefully further datasets that will become available in the future.

2. THE WHS DATASET

WHS define 53 surface type classifications and each of these classifications is divided into 24 basic vegetation components from which various parameters are calculated. For each grid box there is defined a primary type, assumed to occupy 75% of the area, and a secondary type which is assumed to occupy the remainder. Full details on the WHS dataset and how it is used will be found in UM Documentation Paper No. 70.

The 1°x1° resolution of the WHS data gives a very poor representation of the surface type over the UK. Figure 1 shows the primary and secondary classifications over the UK. It can be seen that only three classes are defined: arable, meadow and rough grazing. There are about forty model grid points in each of the WHS grid squares (on the current model resolution) so not even features of several grid lengths are being resolved. Thus there is a clear need to obtain data at a higher resolution.

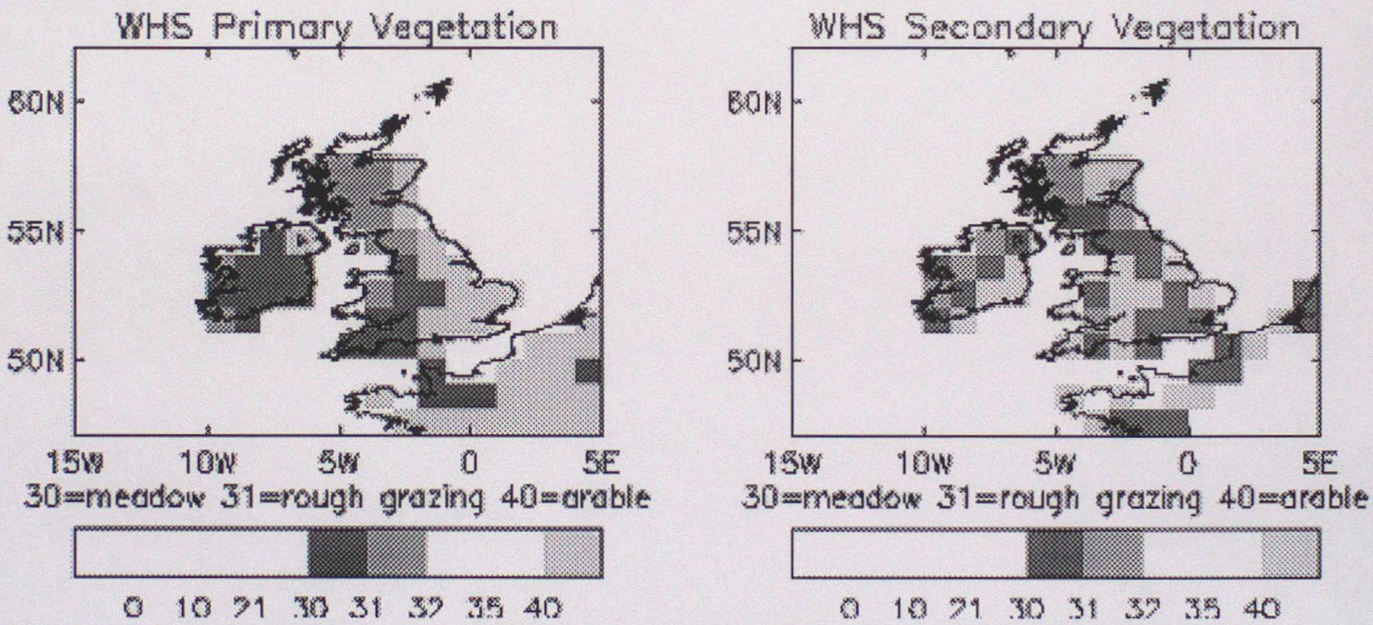

Figure 1: Surface type classifications according to WHS for the UK. The left chart is the primary type and the right chart the secondary type. Dark grey is meadow, medium grey is rough grazing and light grey is arable.

3. ITE NATLAC DATASET

The Institute of Terrestrial Ecology (ITE) has compiled a dataset of land and land use attributes using map and statistical sources for 2858 10x10km grid squares of the Ordnance Survey National Grid which contain land in Great Britain as shown in figure 2. In total there are 126 attributes grouped in categories of physiography, climate, geology, soils, topography, land use, agricultural land classification and conservation status.

Figure 2: The 2858 grid squares in the ITE dataset.

For our purposes we are only interested in the land use attributes with some additional information taken from the physiography and topography attributes.

Land Use Attributes

There are 22 attributes in total:

Agricultural land use (% of grid square)

1.Farmland under cultivation for crops other than grass.
2.Farmland under sown ("improved") grass.
3.Farmland under rough grazing.
4.Total farmland (sum of 1-3).

Other major land uses (% of grid square)

5.Forest and woodland.
6.Urban Development.
7.Balance of land in uses other than those stated above.

Individual agricultural crops (% of grid square)

8.Wheat.
9.Barley.
10.Oats.
11.Potatoes.
12.Sugar Beet.
13.Stockfeed crops.
14.Horticultural crops - vegetables.
15.Horticultural crops - fruit.
16.Total of crops listed above.

The remaining attributes are livestock numbers and are disregarded.

The amount of forest and woodland and urban were taken from maps and all other attributes are taken from official agricultural statistics. These statistics are collected annually from farmers by the Ministry of Agriculture, Fisheries and Food in England and Wales and by the Department of Agriculture and Fisheries in Scotland.

The dataset uses statistics from 1972 so it may be considered rather dated. However, the broadscale distributions of the major attributes have not changed significantly.

The total attributes in any single grid square should equal the total amount of land in that grid square, generally 100% but coastal squares are included. However, for many squares this was found not to be the case. The reasons for this appear to be in the basic data from which the dataset was compiled. Agricultural information is compiled on a parish by parish basis but for farms that extend over more than one parish the data tends to be included only for the parish in which the farmhouse is located. The parish information then needs to be mapped to the grid squares and this has been done by estimating the fraction of the parish in each grid square. There is also a slight discrepancy between statistics collected in England and Wales to those collected in Scotland due to different land type categories used.

These difficulties mean that the quantitative detail on a grid square by grid square basis should be treated with some caution. However, it is felt that the regional detail is far more reliable.

Full details of the NATLAC dataset may be found in Ball et al (1983).

Physiological and Topographical Attributes

The physiological attributes of percentage of land and percentage of inland water (lakes) and the topographical attributes of amount of high ground were used in the processing of the land use attributes.

4. PROCESSING THE ITE DATASET

This section describes the processing that was performed to the ITE dataset so that it could be used in the UM system.

The ITE data is on the Ordnance Survey National Coordinate grid in 10km grid squares. However, for the UM data is more conveniently represented on a latitude-longitude grid. Therefore, the ITE data was transferred to a regular 10'x10' latitude-longitude grid with the NW corner at 61N 351E. Each ITE grid square was mapped to the nearest grid point and for those grid points which had more than one ITE grid square mapped to it an average value was taken. This resolution was chosen as it gave the most complete coverage and it is believed that the averaging employed compensates for some of the errors in the basic data as described above.

By definition all grid squares in the ITE dataset contain at least some land. However, those grid squares which contain only a small fraction of land were disregarded. Only those grid squares which contain at least 50% land, as given in the physiographic dataset, were used.

Nine classes of land use were defined as listed below:

1: Arable cropland
2: Nursery and market gardening
3: Improved grassland
4: Rough grazing
5: Coniferous woodland
6: Deciduous woodland
7: Urban
8: Inland water
9: Pasture and shrub

These classes were determined from the ITE classifications as detailed below.

Arable Cropland

This is all cultivated land (ITE use 1) minus potatoes (ITE use 11) and horticulture (ITE uses 14 and 15). It was considered not worthwhile in trying to differentiate between crops. Even though the main areas of arable crop growing have not changed, the distribution of crops within them would have changed substantially. New crops have been introduced such as oil seed rape and some land has been put out of production as part of the 'set aside' policy. Furthermore, the physical properties of different crops are very similar and too small to be represented even on the scale of the Mesoscale Model.

Nursery and Market Gardening

This is the sum of all horticulture (ITE uses 14 and 15) and potatoes (ITE use 11).

Improved Grassland
This is ITE use 2.

Rough Grazing
This is ITE use 3.

Coniferous and Deciduous Woodland
The ITE use 5 is all woodland, both coniferous and deciduous, but the UM ideally requires separate knowledge of coniferous and deciduous woodland. Therefore, the percentage of woodland was subjectively divided between coniferous and deciduous woodland.

Woodland below 10% cover was assumed to be all deciduous woodland, representing isolated copses etc and the remnants of the traditional British forest. Secondly, all woodland over high ground, defined as 50% of the grid square above 800' as given in the topographical dataset, was assumed to be 100% coniferous representing forestry plantations. Thirdly, an area of forestry representing Thetford Forest in East Anglia was also assumed to be 100% coniferous. Elsewhere the woodland was divided into 40% coniferous and 60% deciduous except over an area of Berkshire, North Hampshire and Surrey which was classed as 80% coniferous and 20% as there are substantial areas of forestry plantations in this area.

All coniferous forests are further assumed to consist entirely of evergreen trees and densely planted and all deciduous forests are further assumed to consist entirely of broad-leaved trees and rather more open.

Urban
This is ITE use 6.

Inland Water
This is taken from the physiographic dataset.

Pasture and Shrub
This is an additional class introduced so that in each grid square the total land use equalled 100% (coastal squares are assumed to consist entirely of land, ie fractions of sea are ignored). The sum of the previous eight classes was found and the shortfall calculated. This shortfall was then assumed to be rough grazing for high ground, defined as above, and coastal squares, defined as having a fraction of land between 50% and 99%. The reasoning is that both high ground and coastal squares are exposed so rough grazing is more appropriate. For the remaining squares the shortfall was defined to be pasture and shrub.

Some squares were found to have to the total land type exceeding 100%. In these squares each type was reduced by 1%, starting with the dominant type, until a total of 100% was obtained.

The result is a dataset giving the percentage of each of the above nine classes on a 10'x10' grid as shown in the following figures.

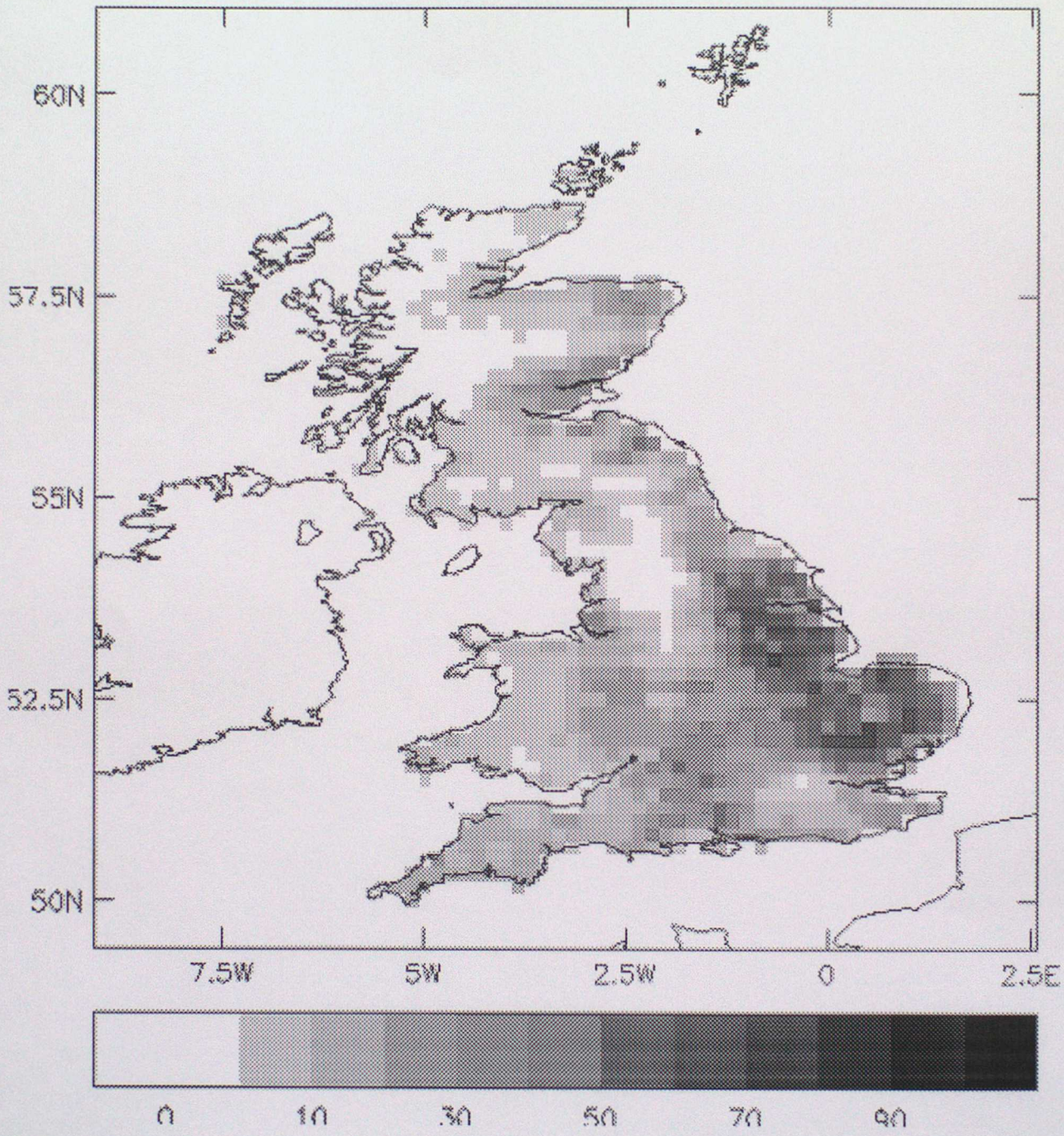

Figure 3: Percentage of arable cropland.

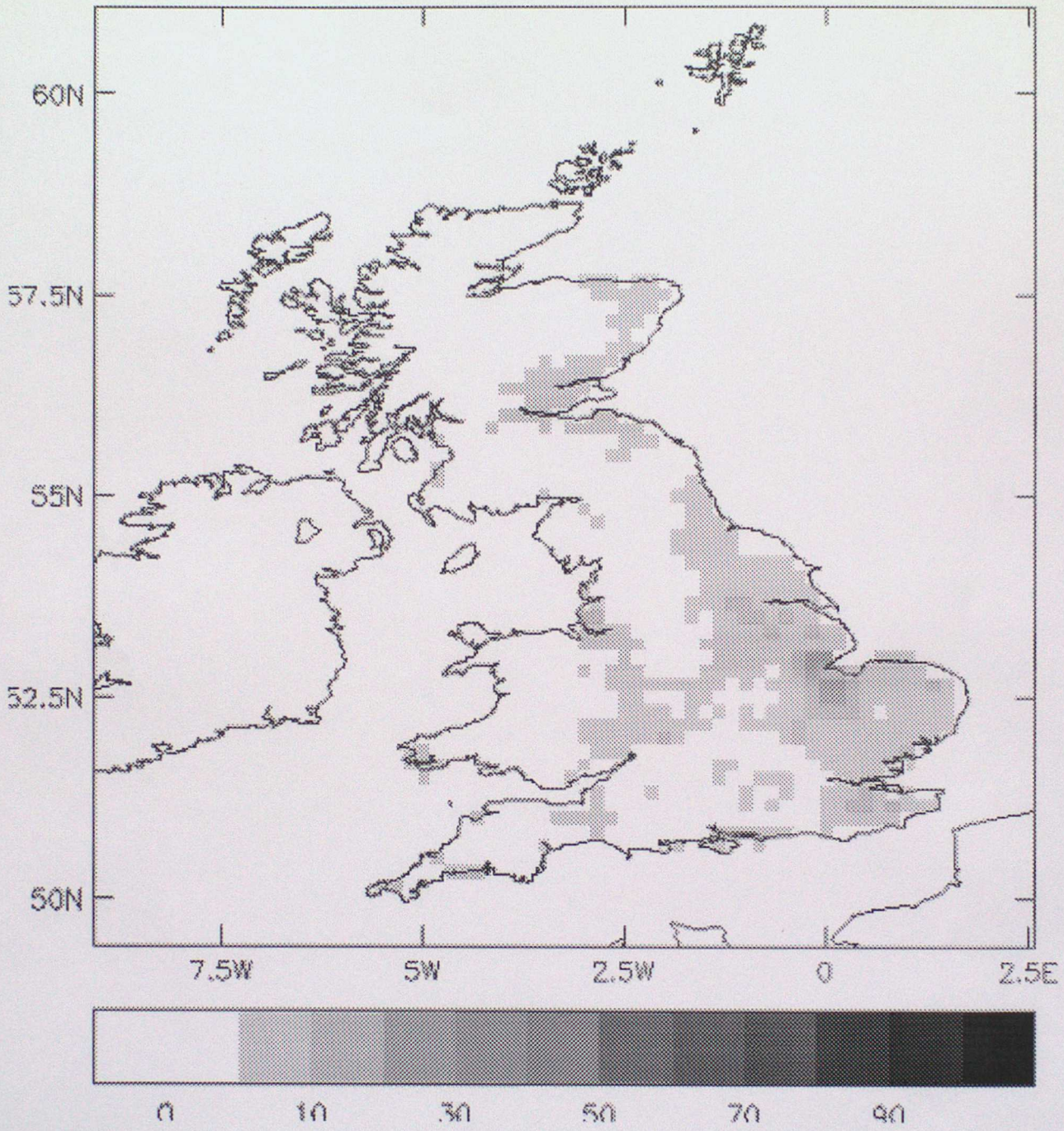

Figure 4: Percentage of nursery and market gardening.

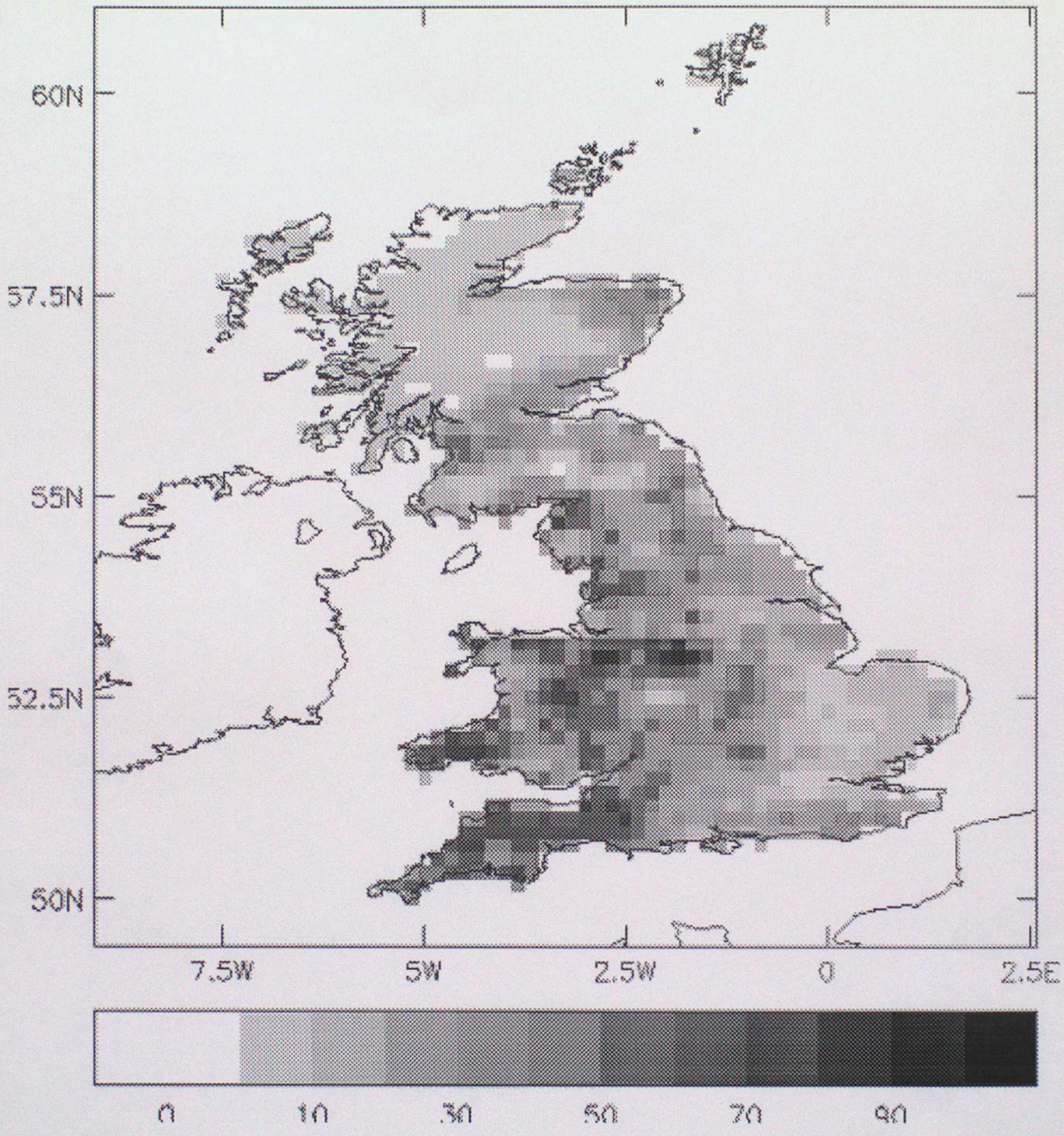

Figure 5: Percentage of improved grassland.

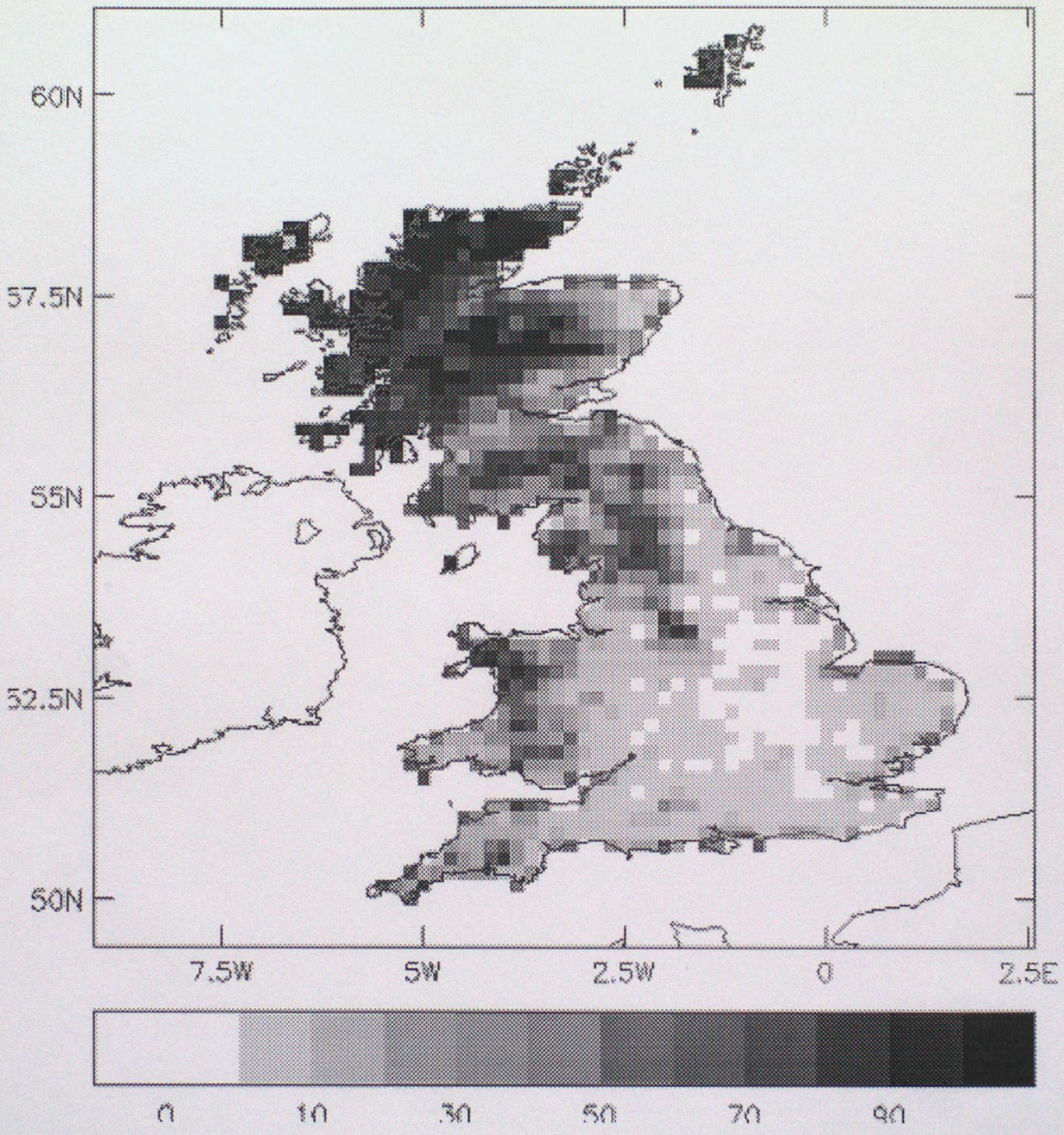

Figure 6: Percentage of rough grazing.

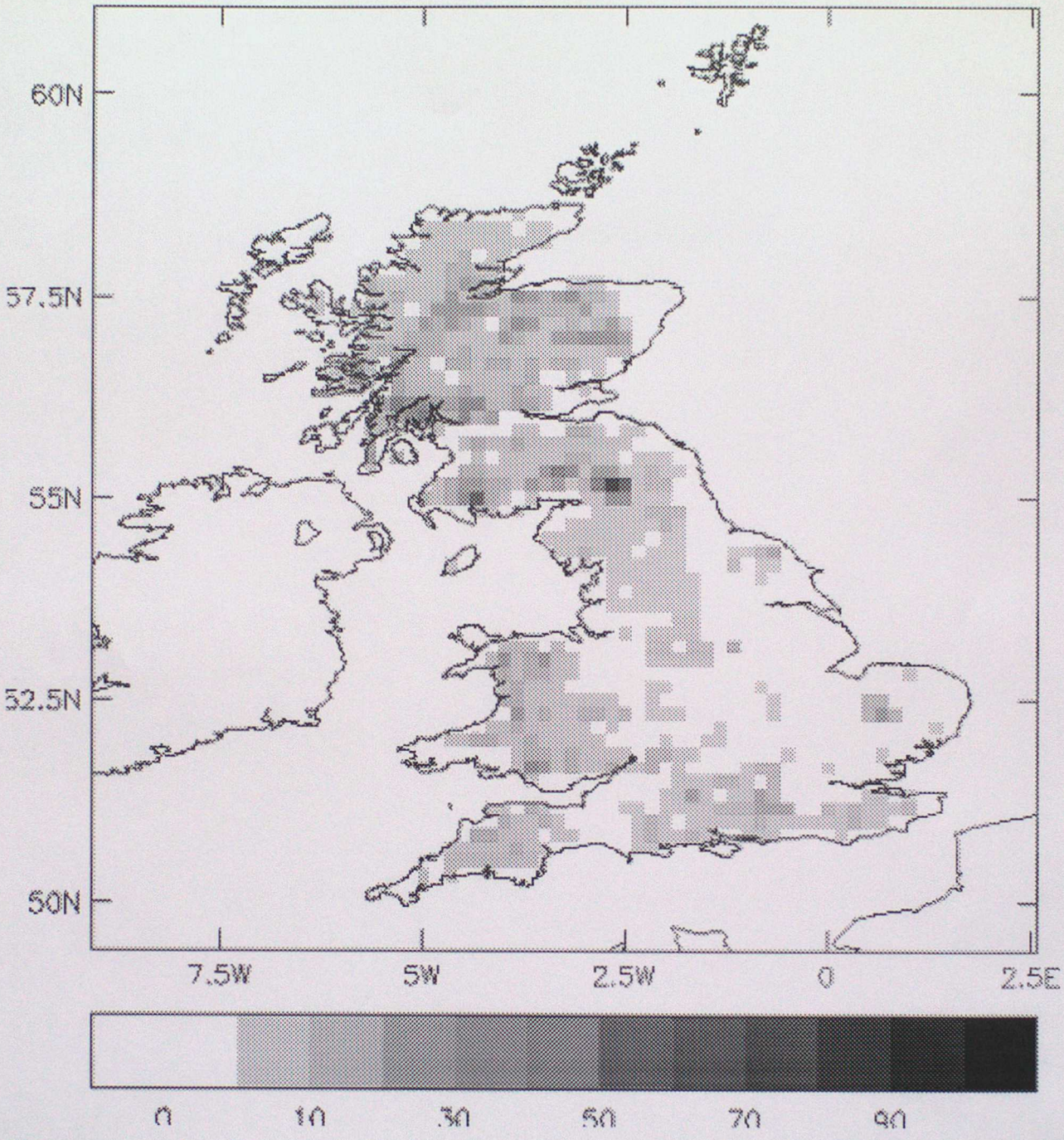

Figure 7: Percentage of coniferous woodland.

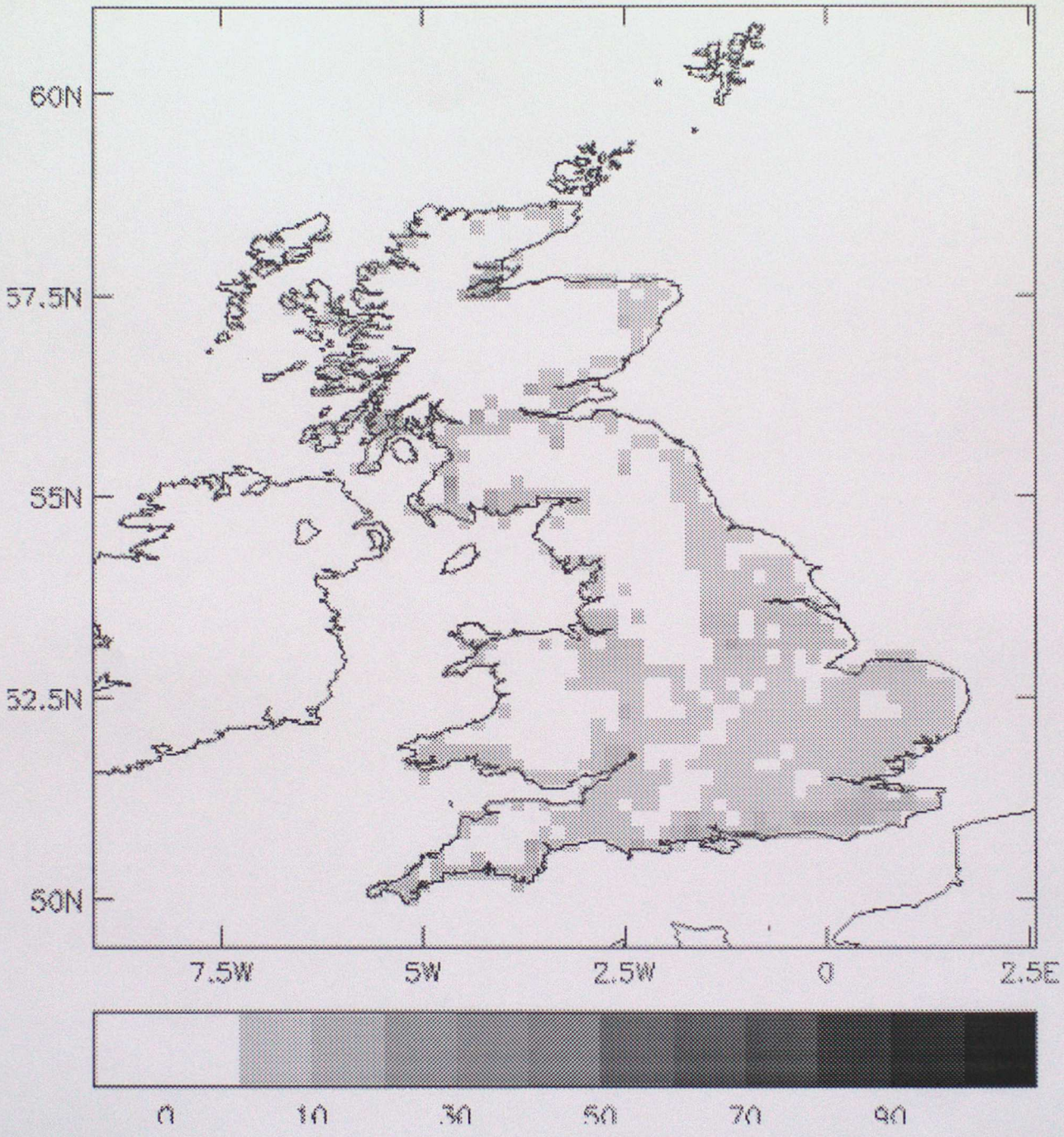

Figure 8: Percentage of deciduous woodland.

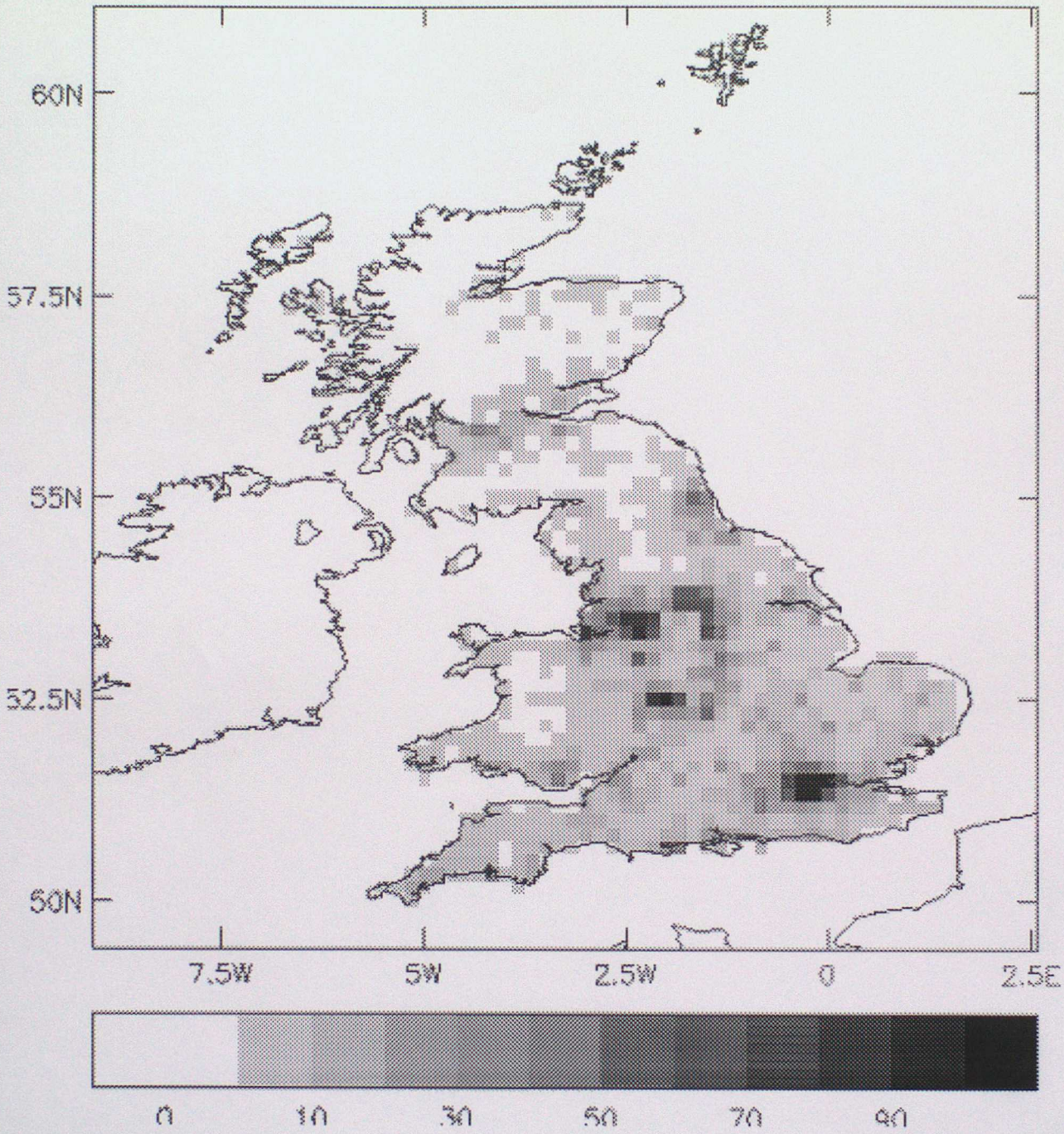

Figure 9: Percentage of urban.

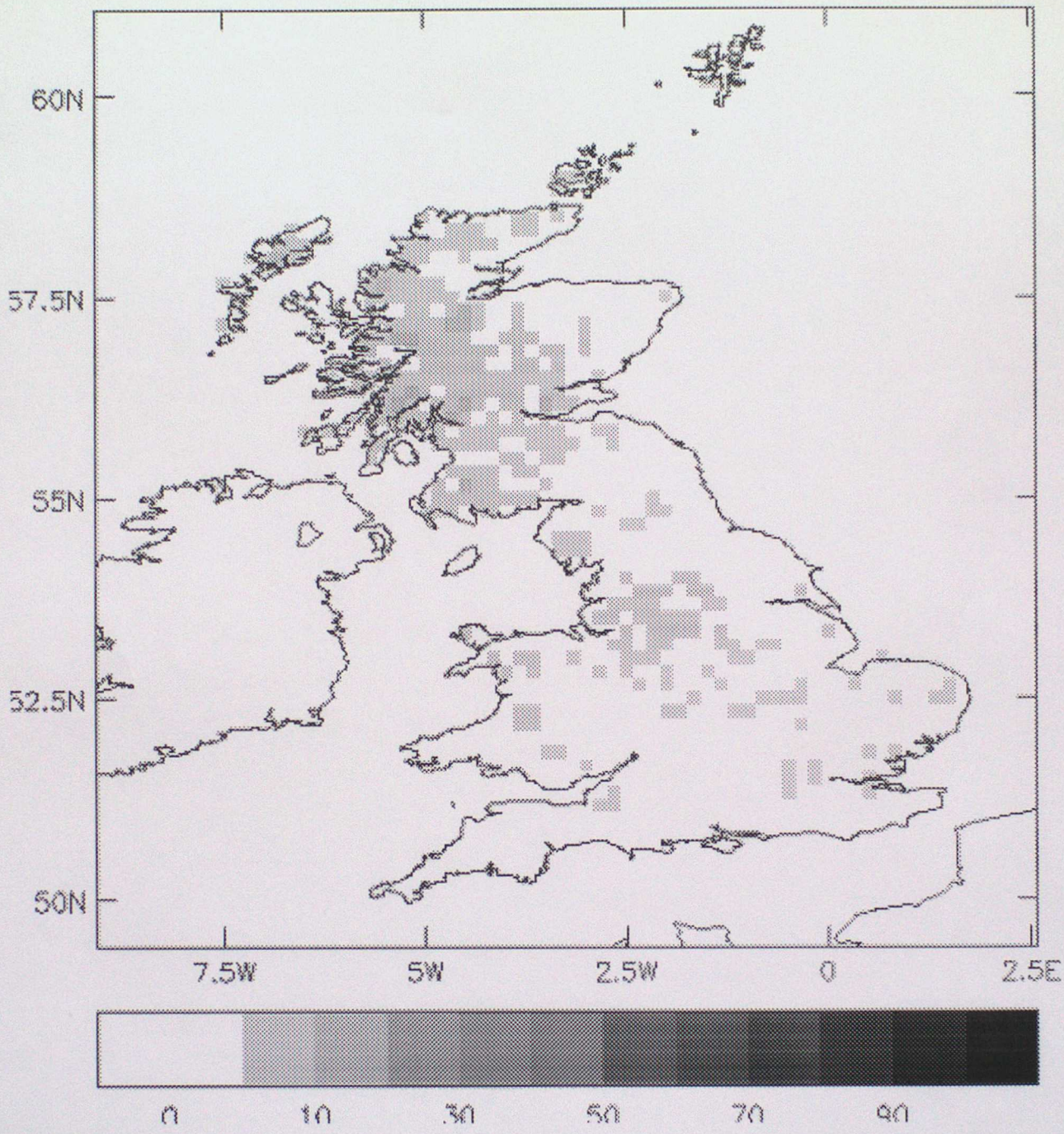

Figure 10: Percentage of inland water.

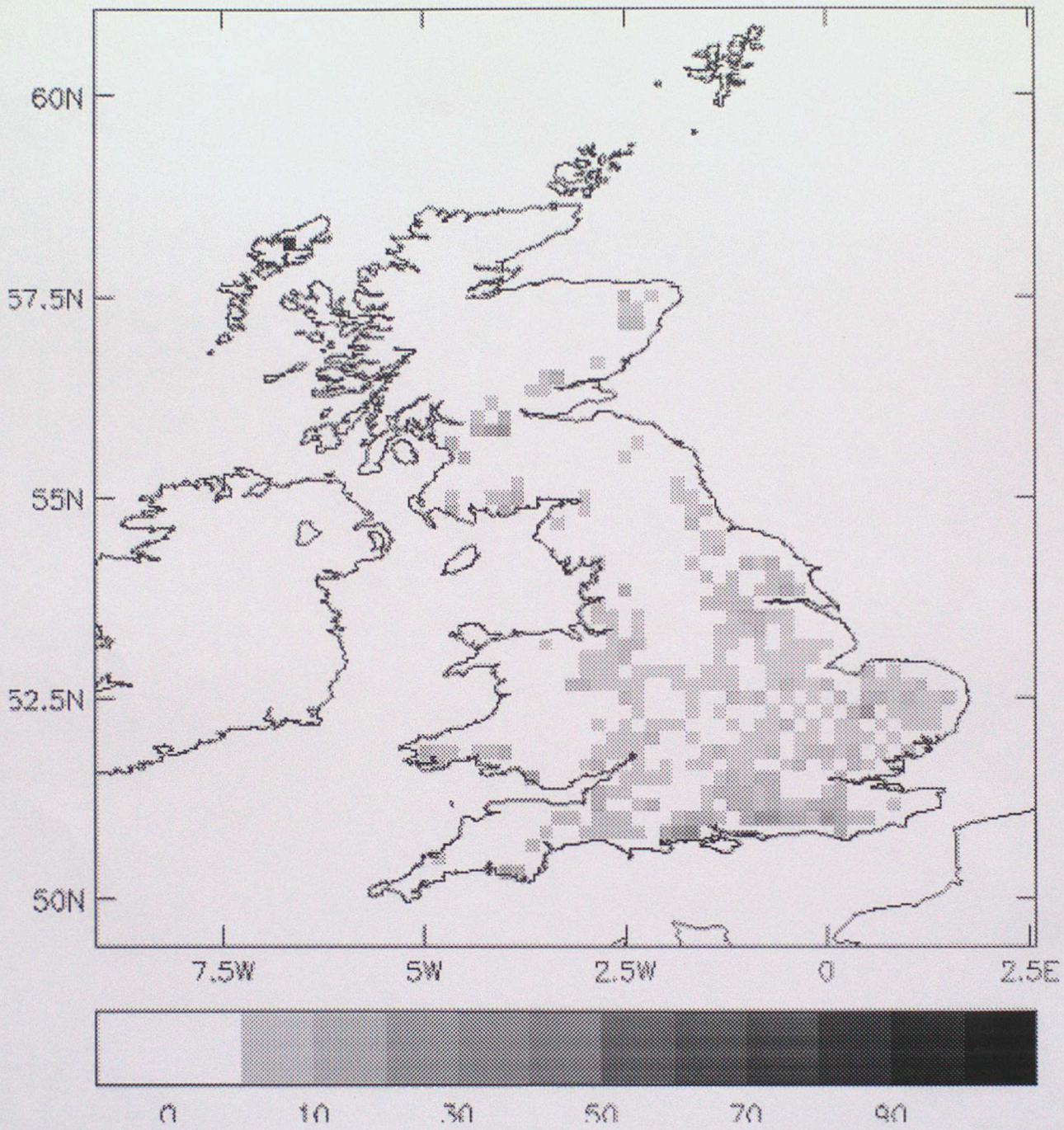

Figure 11: Percentage of pasture and shrub.

From these maps the general distribution of the different land use types in England, Scotland and Wales can be seen. Arable cropland is dominant on the eastern side of England whilst Scotland and northern England is dominated by rough grazing which could be considered as moorland on the high ground. The forestry plantations over Scotland and Wales are evident and there is a mixture of land types across southern Britain. The urban centres of London, Birmingham and Manchester are also clearly seen.

5. USING THE ITE DATASET

The land use data is not used directly in the model in the current surface parametrization schemes. Instead it is used to calculate a series of parameters that describe the physical properties of the surface or vegetation. The parameters currently used are: root depth (D_R), snow free albedo (α_O), deep snow albedo (α_S), surface resistance to evaporation (r_S), surface roughness (z_0), soil water infiltration factor (β_V), vegetation fraction (ν) and canopy capacity (C_M). Figure 12 is a schematic diagram illustrating these parameters.

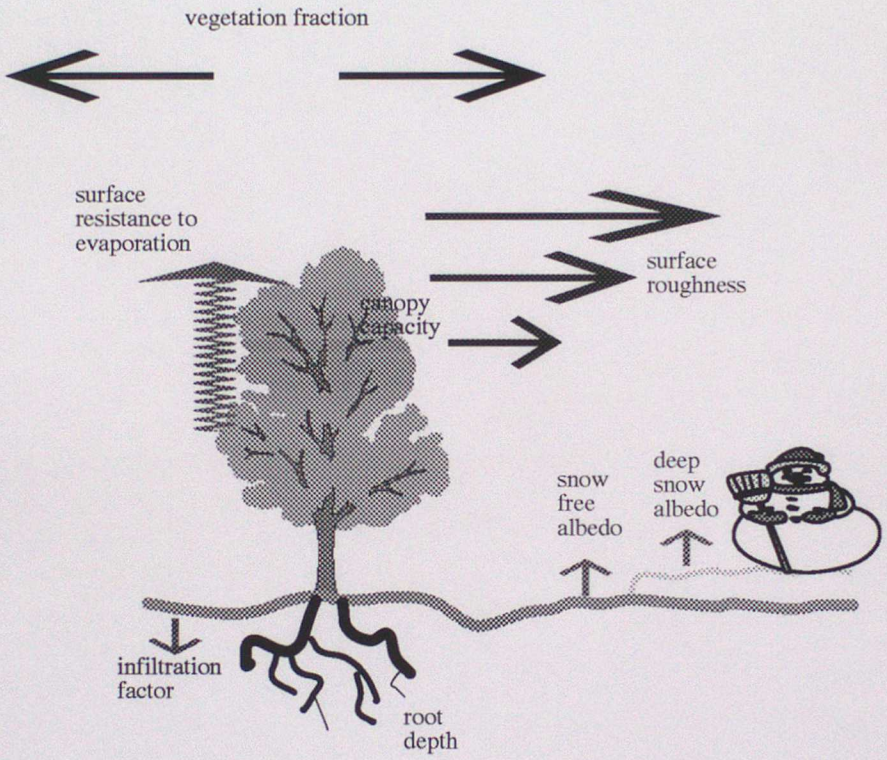

Figure 12: Schematic diagram of parameters that describe the physical properties of the vegetation.

These parameters have been calculated on the same 10'x10' grid as the ITE dataset using an approach based on that employed with the WHS dataset.

As with the WHS dataset, each of the surface type classes is divided into basic components. The reason for doing this is that an arable field for example will not consist entirely of arable crops. There will be areas of bare soil and perhaps grass around the borders. Furthermore, the parameters are fixed throughout the year and the proportions of the basic components attempt to represent an annual mean.

Nine basic components are required for the ITE data:

1: Inland water
2: Evergreen needle-leaf tress
3: Deciduous broadleaf trees

4: Deciduous shrub
5: Short grass and forbes
6: Long grass
7: Arable
8: Urban
9: Soil

Table 1 gives the annual mean percentage of each of the basic components for each of the surface type classes.

	Basic Component								
ITE class	1	2	3	4	5	6	7	8	9
arable					20		60		20
nursery							90		10
grassland					90				10
grazing				10	70	5			15
coniferous		90			5				5
deciduous			60	5	30				5
urban				20				70	10
water	100								
pasture				15	75				10

Table 1: Annual mean percentage of each component in each of the surface type classes.

Values for each of the required vegetation parameters for each of the basic components have been obtained from the scientific literature sources as detailed in UMDP 70. These values are shown in table 2.

	Vegetation Parameters							
Basic Component	D_R (m)	α_o	r_s (sm^{-1})	z_0 (m)	C_{M_2} (kgm^{-2})	ν	β_v	α_s
1	1.00	0.06	0.0	0.0003	0.00	0.00	1.00	0.80
2	0.90	0.14	85.0	1.0000	1.20	1.00	6.00	0.20
3	1.20	0.13	100.0	1.0000	0.60	1.00	6.00	0.30
4	0.90	0.16	80.0	0.4000	1.00	1.00	3.00	0.50
5	0.50	0.19	60.0	0.0100	0.50	1.00	1.50	0.70
6	0.70	0.20	80.0	0.0400	0.70	1.00	2.00	0.60
7	0.70	0.20	60.0	0.0400	0.50	1.00	2.00	0.80
8	0.00	0.18	200.0	1.5000	0.50	1.00	0.10	0.40
9	0.10	0.17	100.0	0.0003	0.50	0.00	0.50	0.80

Table 2: Values for each of the required vegetation parameters for each of the basic vegetation components.

For each surface type class j, an annual mean parameter value, P_j, is calculated by multiplying the parameter value for each basic component i, P_i, by the fraction of that component, α_i. This is represented by the equation:

$$P_j = \sum_{i=1}^{9} \alpha_i P_i / 100$$

Then for each grid box k a mean value of the parameter, P_k, is calculated by multiplying each P_j by the fraction of that class in the grid box, δ_j.

$$P_k = \sum_{j=1}^{9} \delta_j P_j / 100$$

The result is each of the eight required parameters on the same 10'x10' latitude-longitude grid as the original data. The parameters then need to be transferred to the grid used in the Mesoscale Model. In order to preserve as much detail as possible it was decided to use an area averaging interpolation technique. However, it was found to be not possible to interpolate directly from the ITE grid to the model grid using this method as the grids are on different orientations and the model grid covers a much wider area. Therefore, a two step interpolation approach was adopted. The parameters are firstly transferred to a grid lying entirely within the ITE domain but at the resolution of the Mesoscale Model using area weighted interpolation. Data from this grid is then interpolated to the Mesoscale Model grid using a bi-linear technique, the Mesoscale coordinates first being unrotated. The model area lying outside the ITE domain still uses data derived from the WHS dataset.

The following figures show the eight parameters on the Mesoscale Model grid.

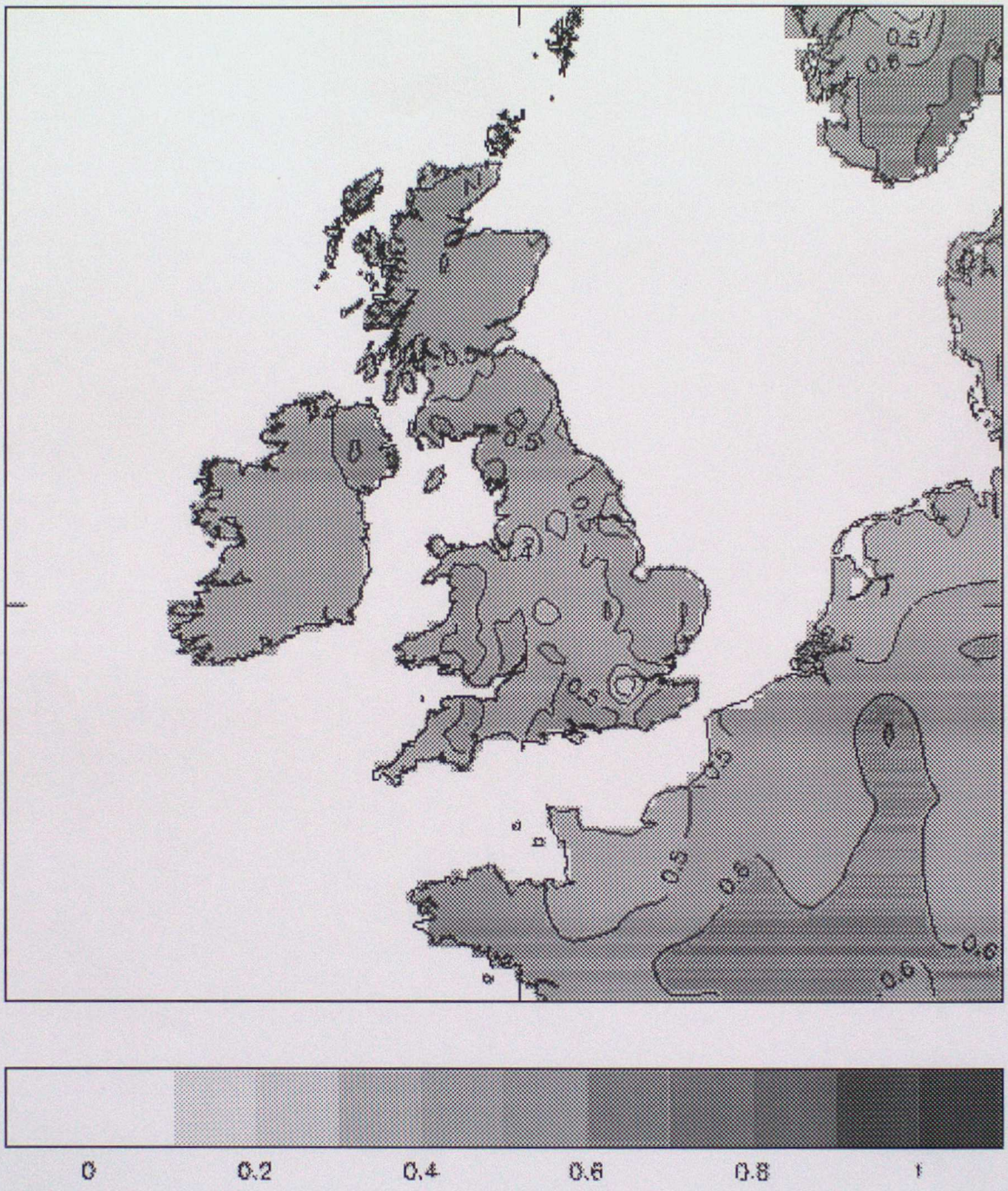

Figure 13: Root depth parameter (m) .

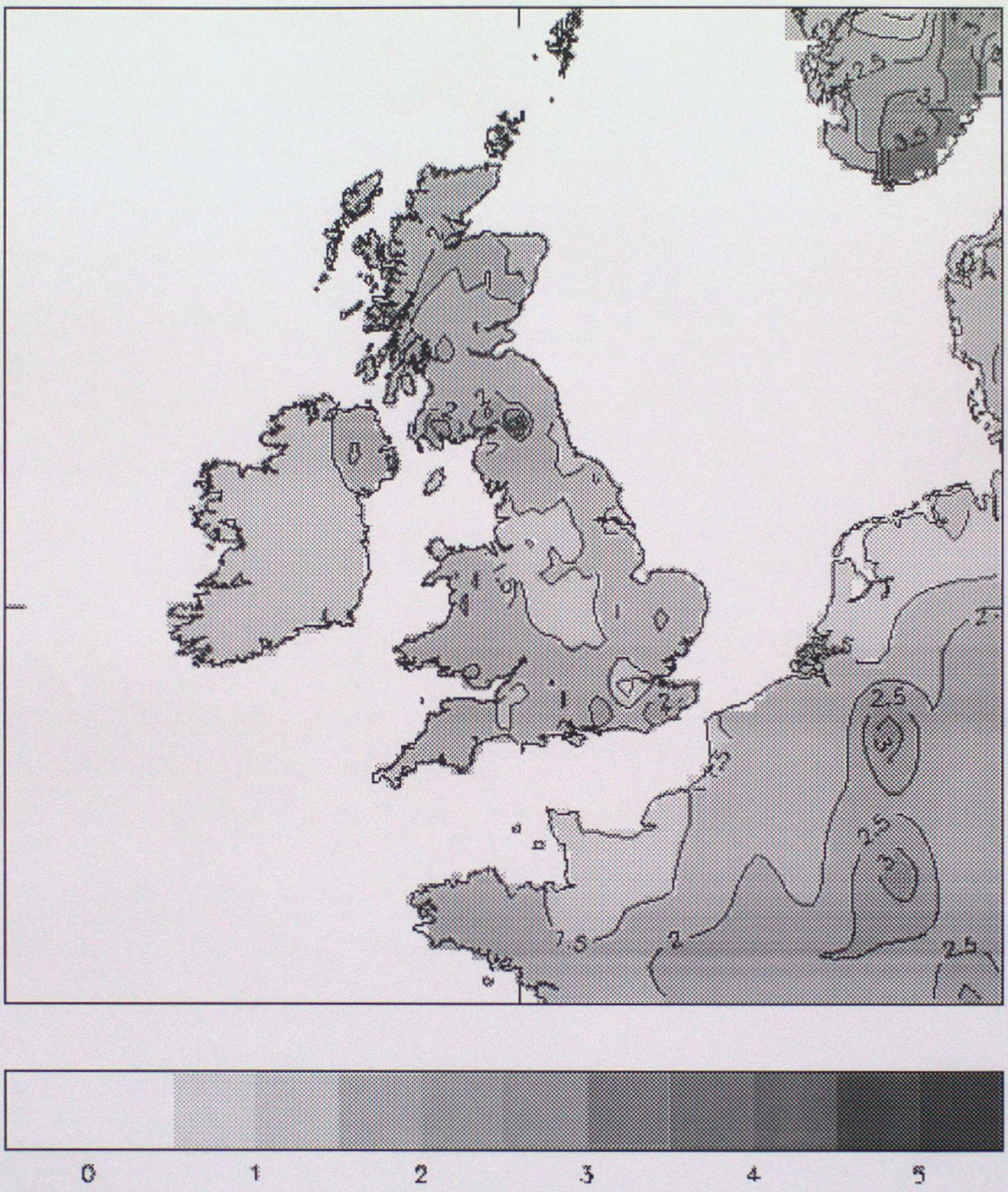

Figure 14: Soil water infiltration factor parameter.

Figure 15: Surface resistance to evaporation parameter (sm^{-1})

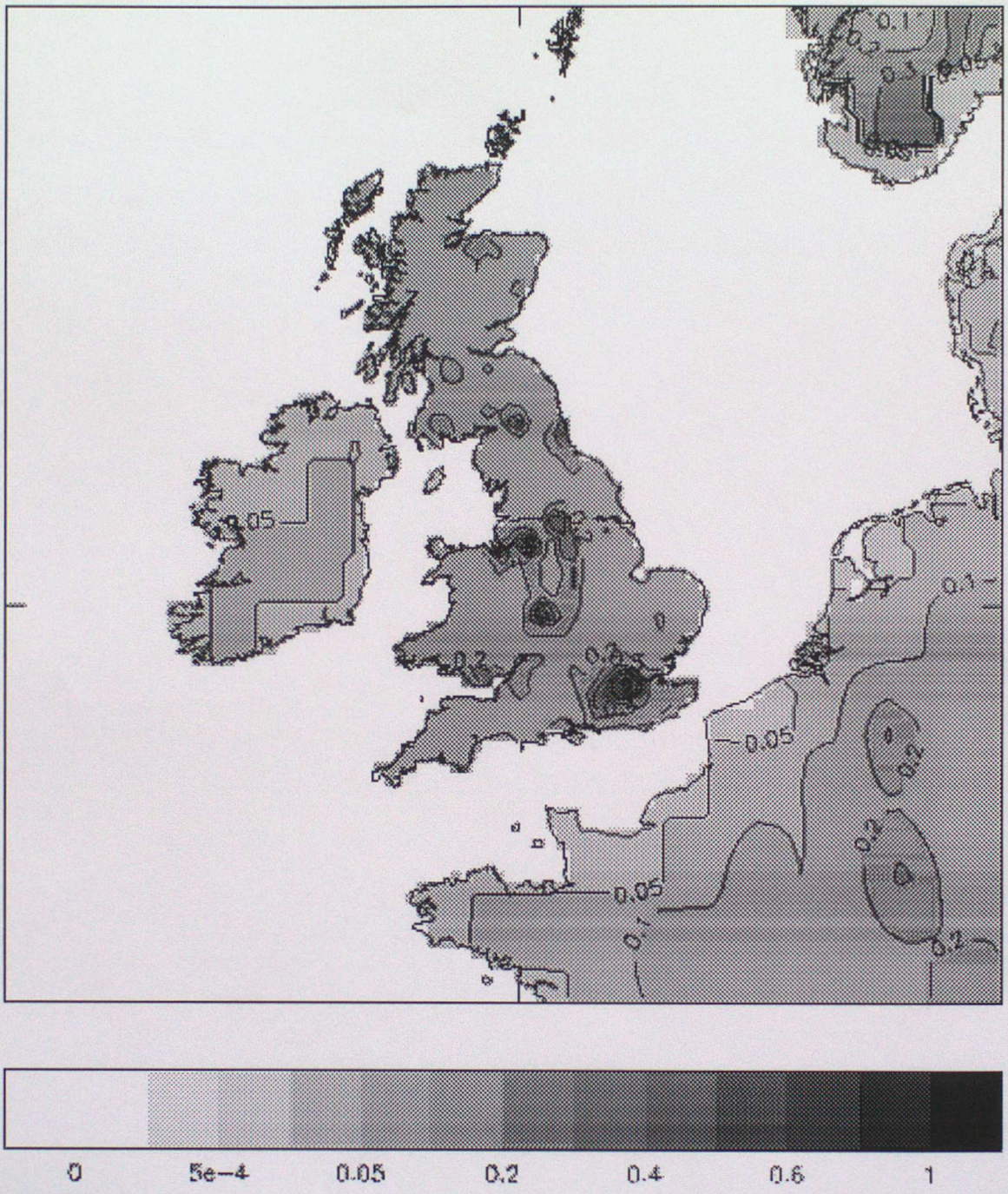

Figure 16: Vegetation roughness parameter (m).

Figure 17: Canopy capacity parameter (kgm^{-2}).

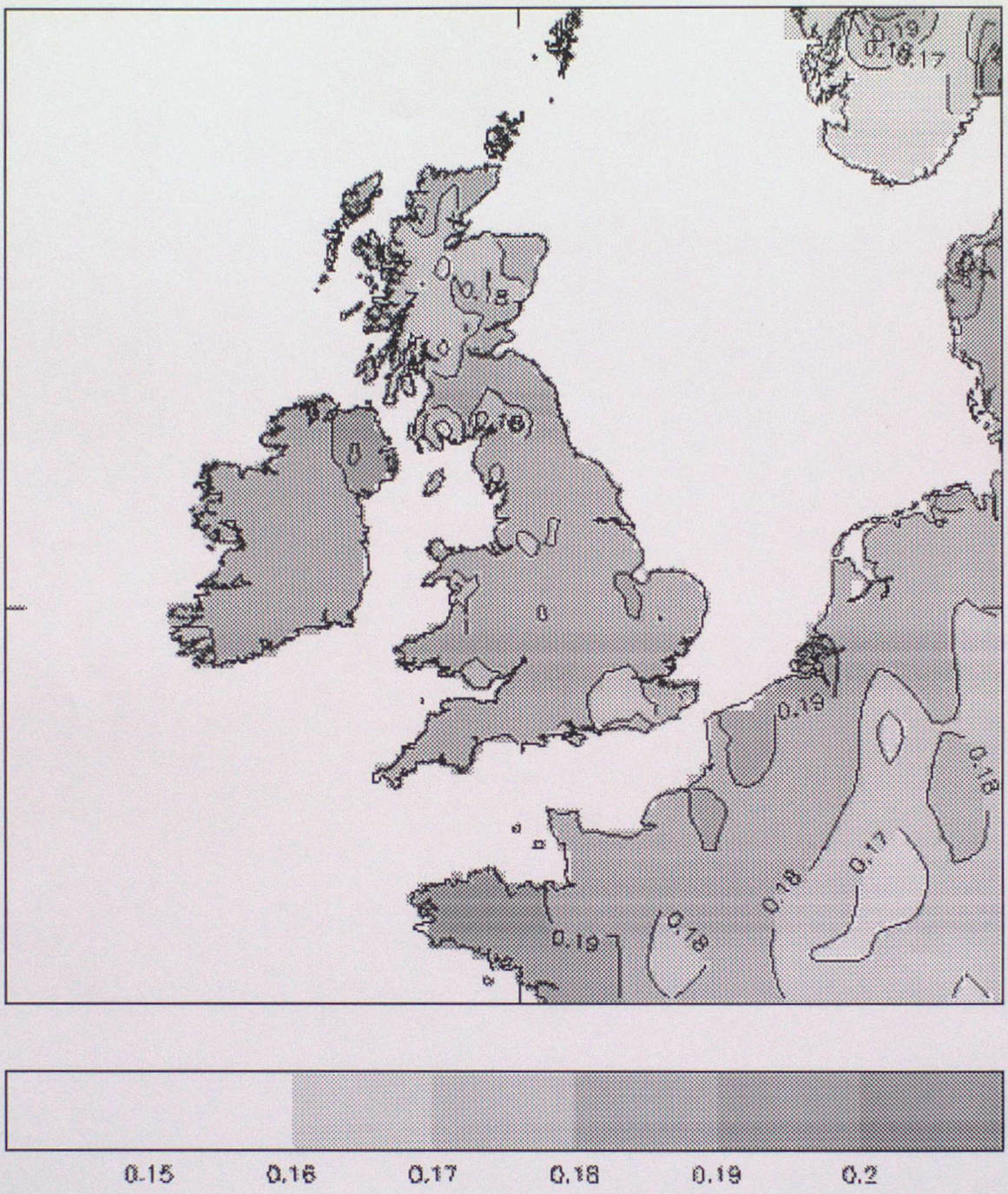

Figure 18: Snow free albedo parameter.

Figure 19: Deep snow albedo parameter.

Figure 20: Vegetation fraction parameter.

6. CASE STUDIES

The new parameter data fields were tested in the Mesoscale Model to determine what, if any, impact there was. The cases were chosen to represent a range of meteorological conditions.

CASE 1: DT 00Z 17 January 1995. Strong winds and a front crossing the UK driven by a vigorous depression to the north west.

CASE 2: DT 00Z 28 June 1995. Clear, warm summer's day (using MORECS soil moisture in both runs).

CASE 3: DT 18Z 11 October 1995. Widespread radiation fog.

CASE 4: DT 12Z 19 February 1996. Strong winds and widespread, heavy snow across south east.

CASE 5: DT 18Z 27 February 1996. Clear winter's night.

CASE 6: DT 00Z 6 March 1996. Extensive stratus and stratocumulus low cloud.

CASE 7: DT 18Z 16 March 1996. Area of radiation fog over central southern England.

CASE 8: DT 00Z 10 May 1996. Widespread convection.

CASE 9: DT 00Z 19 May 1996. Strong winds and widespread rain.

In addition case 4 was repeated in a 4 assimilation, 2 forecast cycle commencing at 00Z 19 February 1996 with forecasts at 6Z and 12Z.

For each case two experiments were run, control and ITE test. The only difference between the two experiments was that the ITE run included the new vegetation parameters as described above and a new orographic roughness parameter field for reasons explained below.

Version 4.0 of the Unified Model was run with latent heat nudging. Each case included a full 6 hours assimilation including MOPS data.

The largest impact was expected in the 10m wind field due to the change in roughness lengths. Previous experience has shown that the model is fairly sensitive to changes to the roughness field and much work has been done in this area over the past few years such as the introduction of orographic roughness. In fact, in the control run the orographic components over London have been artificially increased.

Some impact in screen temperature and humidity may also be expected due to the different vegetation properties.

Objective Verification

Version 1.7 of the Mesoscale Model station verification system was used to obtain objective verification statistics.

The graphs show verification for screen temperature, 1.5m relative humidity, 10m wind speed and 10m wind vector. The curve starting at DT 00Z is an average of 5 cases and the curve starting at DT 18Z is an average of 3 cases. Also plotted are the single 6Z run and the single 12Z run although less weight should be given to these curves as the single case may not be wholly representative. The control is plotted as a solid line and the ITE run as a dashed line, denoted by the suffix C and I on the legend respectively.

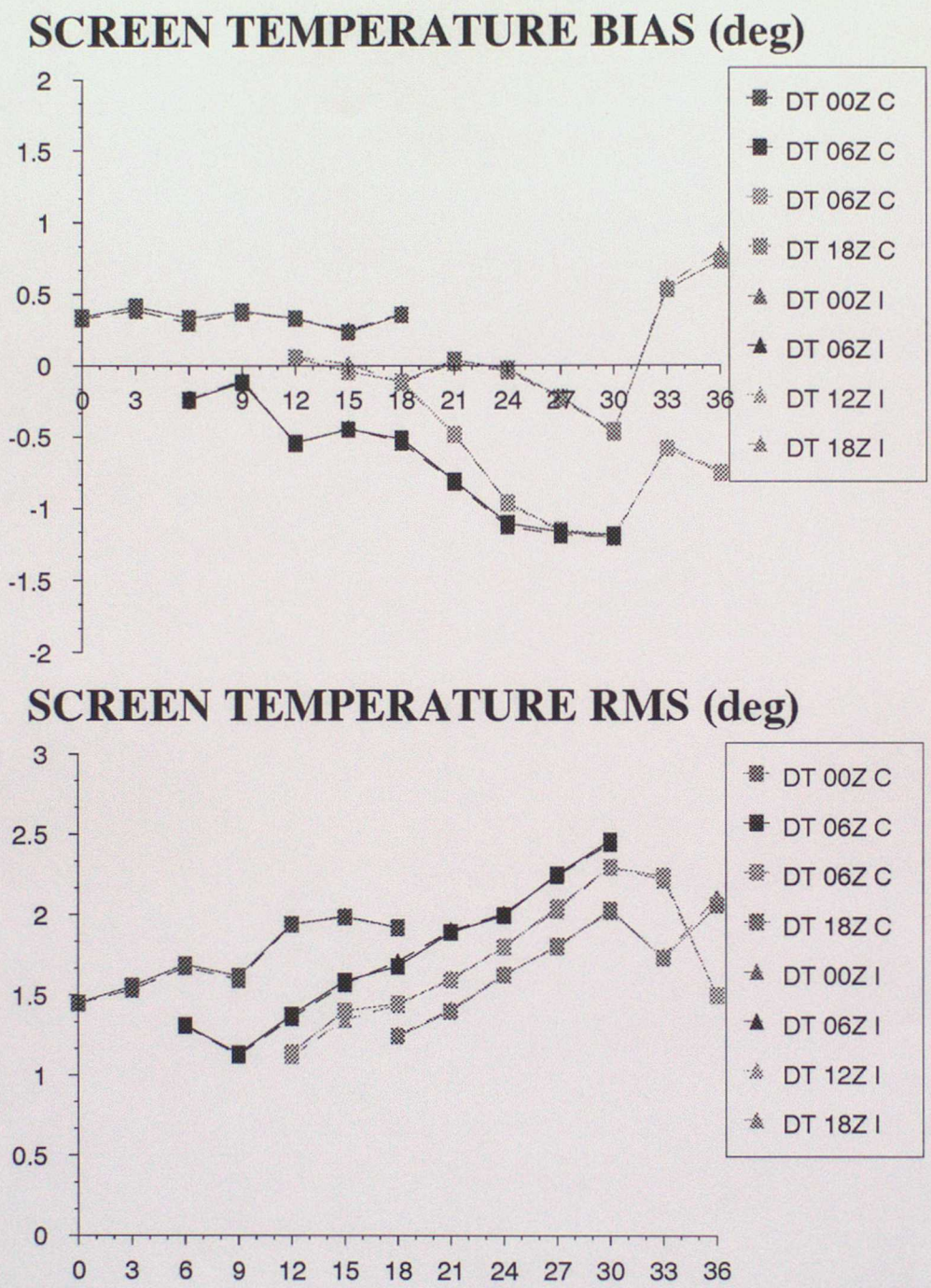

Figure 21: Screen temperature bias and rms errors.

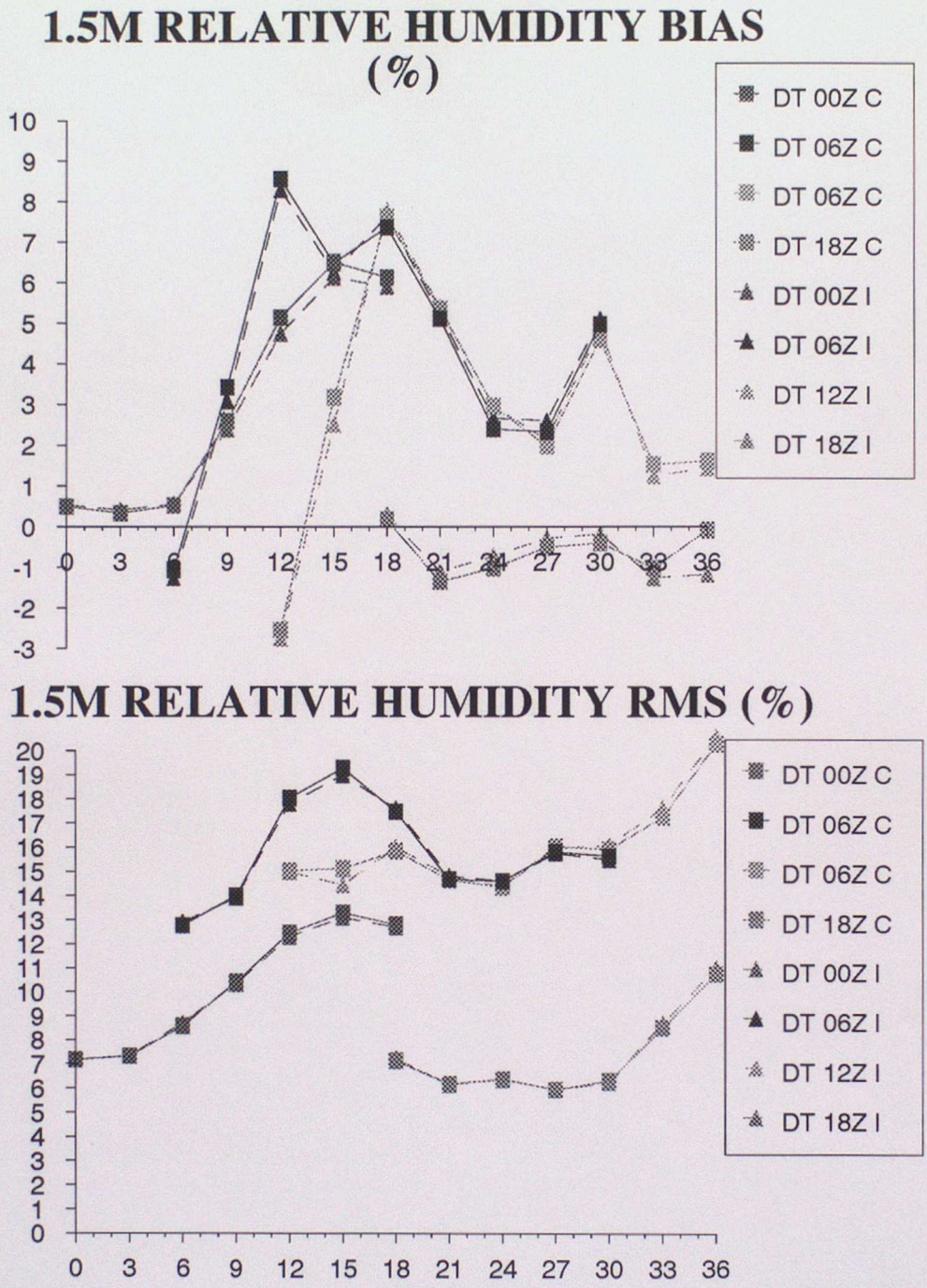

Figure 22: 1.5m relative humidity bias and rms errors.

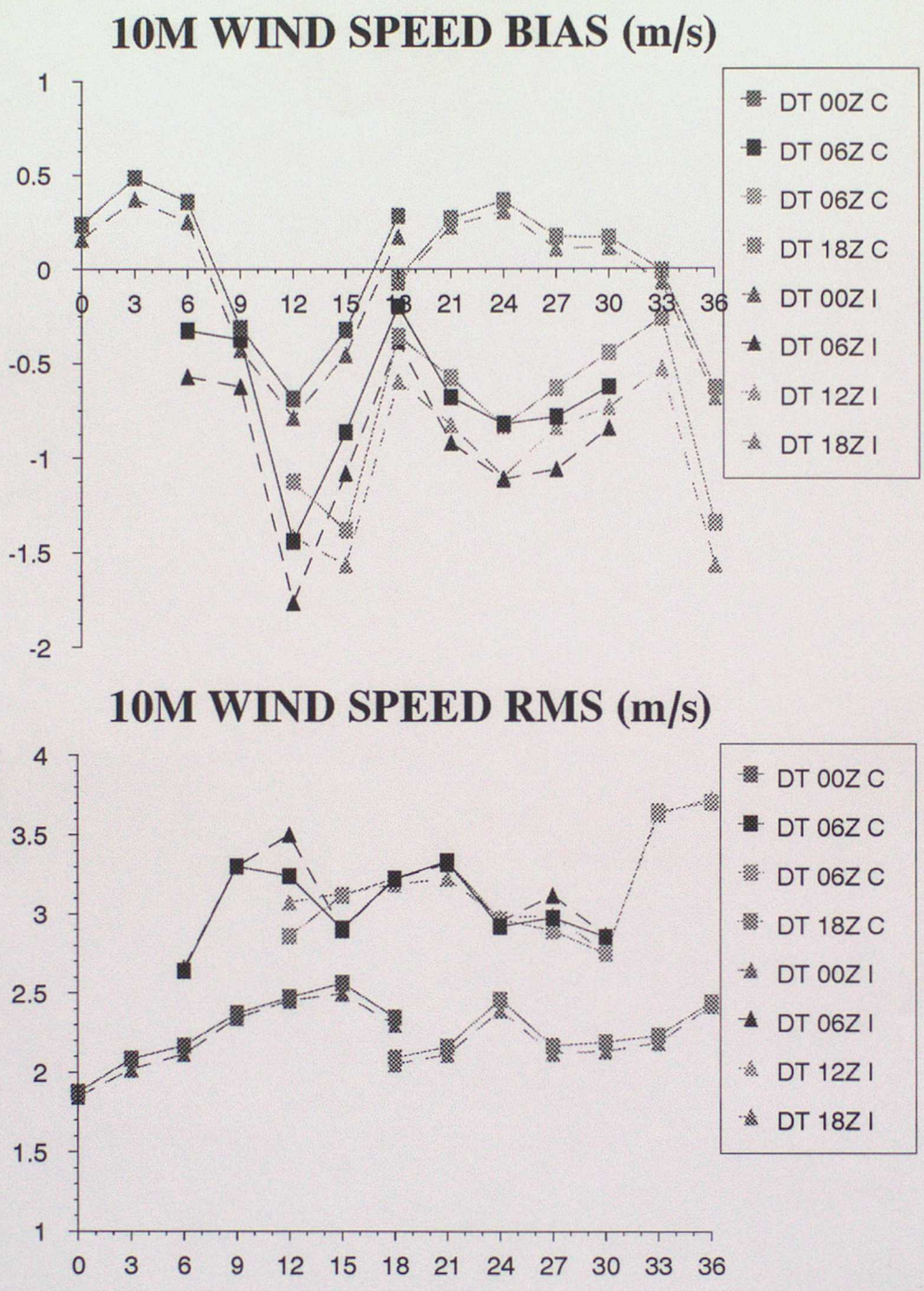

Figure 23: 10m wind speed bias and rms errors.

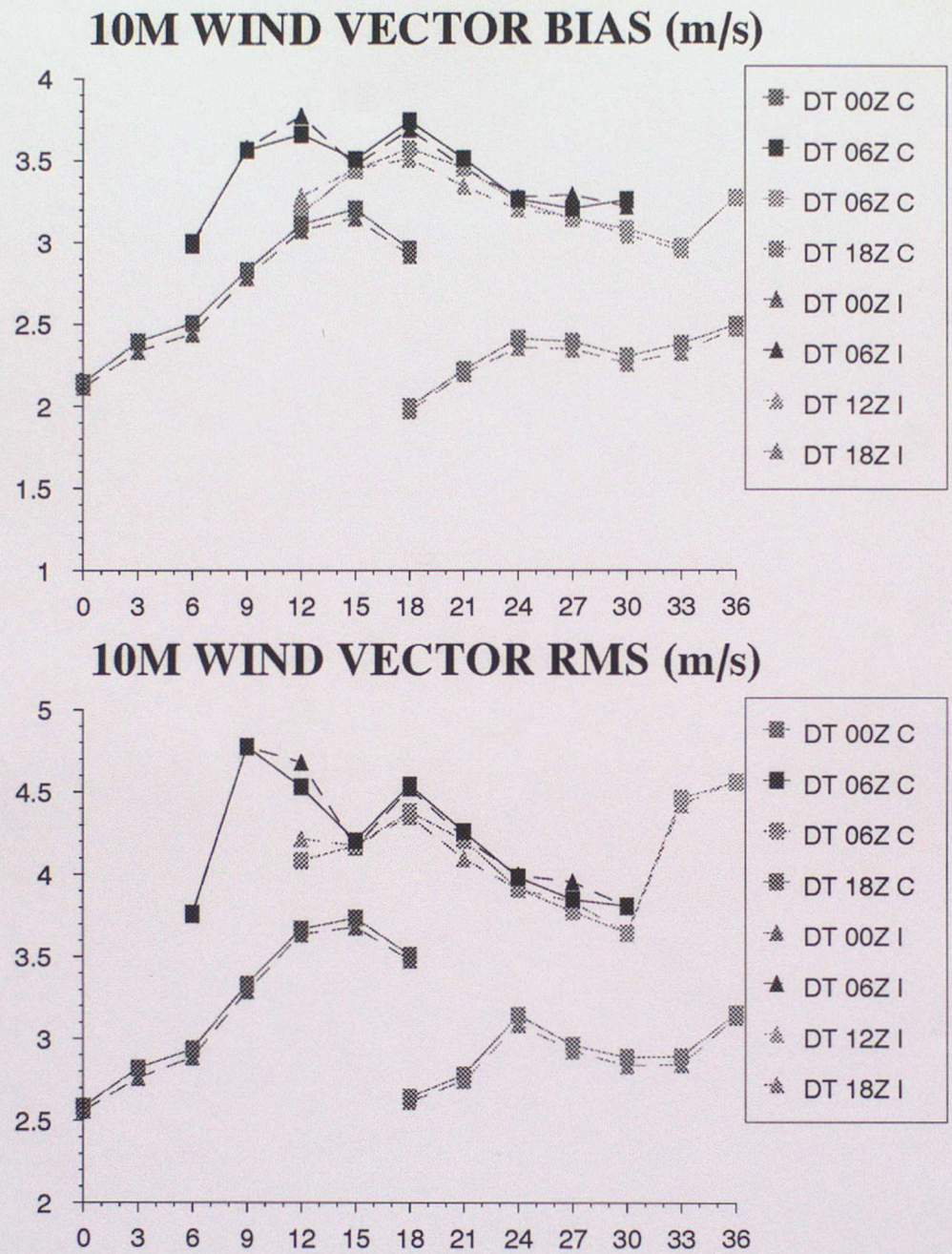

Figure 24: 10m wind vector bias and rms errors.

As expected, most impact is in the 10m wind field. The ITE data has a tendency to decrease the wind speed over England and Wales especially, where the majority of the verifying stations are, and therefore any positive bias is reduced and similarly any negative bias increased, ie made more negative. However, this is not always the case. In the 19 February 1996 case the ITE data actually made a negative bias in the control 10m wind speed slightly less negative. However, the rms errors, which are a better guide to the fit of the data to the observations, show a slight improvement.

There is a much clearer improvement in both bias and rms in the 10m wind vector error which does indicate that the ITE data largely produces a better forecast of the 10m wind. The improvement is about 2%.

Both screen temperature and 1.5m relative humidity show, perhaps surprisingly, little impact. There is a slight improvement in the rh bias in all the runs but little impact in the rms scores and there is very little impact in the screen temperature bias or rms.

The following figures show the verification of 10m wind speed of on the 17th January 1995 case.

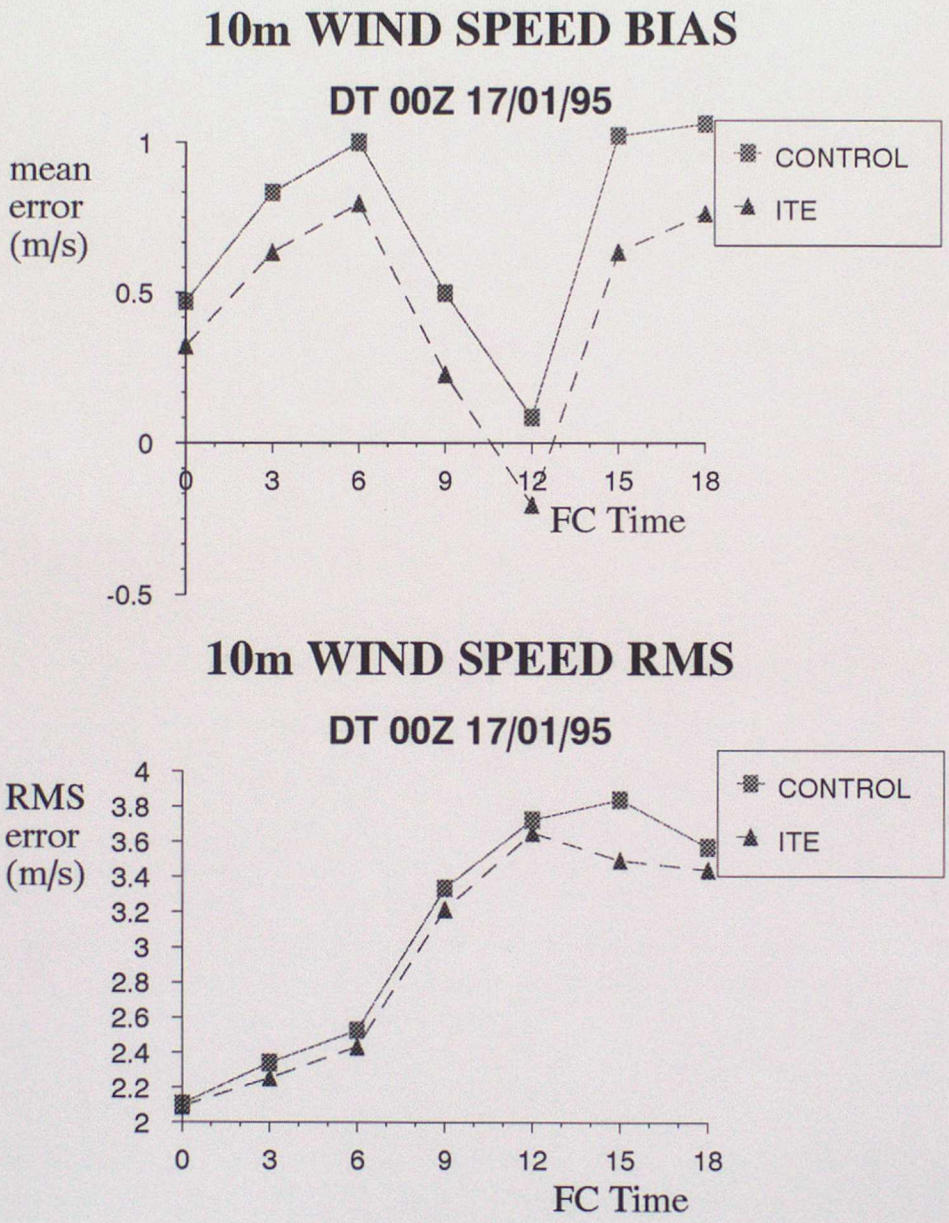

Figure 25: Verification of 10m wind speed for 17th January 1995. Solid line is control, dashed line is ITE data.

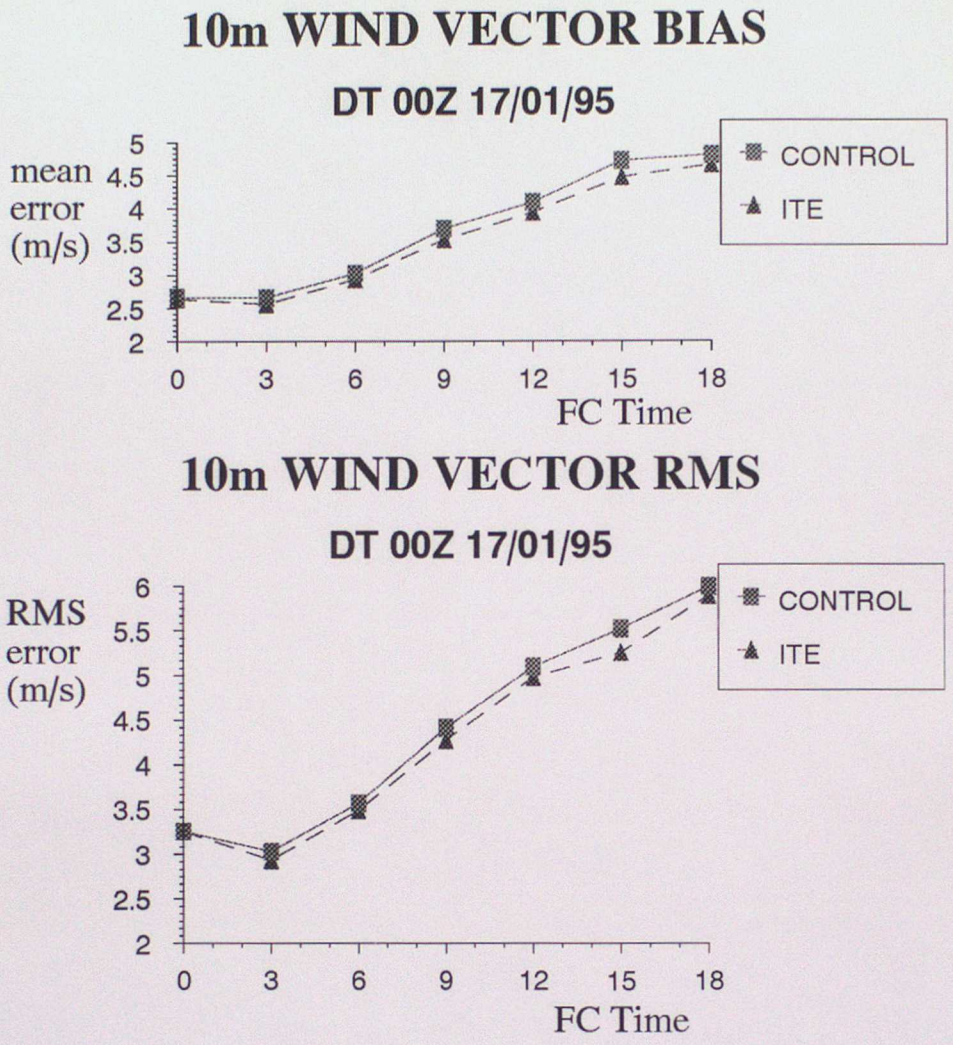

Figure 26: Verification of 10m wind vector for 17th January 1995. Solid line is control, dashed line is ITE data.

There is a clear and consistent improvement throughout the forecast. The reduction in the 10m wind bias is very pronounced, so much so that at T+12 a slight positive bias in control is replaced by a slight negative bias in the ITE run, However, the 10m wind vector bias and rms is reduced in the ITE data indicating that it is a more realistic wind field.

Summary of Objective Verification

The most pronounced effect is in the 10m wind speed with a predominantly positive impact. Other fields show either a slightly positive or neutral impact and certainly no negative impacts were found. The improvement in the 10m wind vector error is about 2% which is a significant improvement.

Subjective Assessment

A number of fields were studied subjectively to try to establish any clear signals the new data is having. Looking at 10m wind speed two signals kept reoccurring and an example is shown in the figures below.

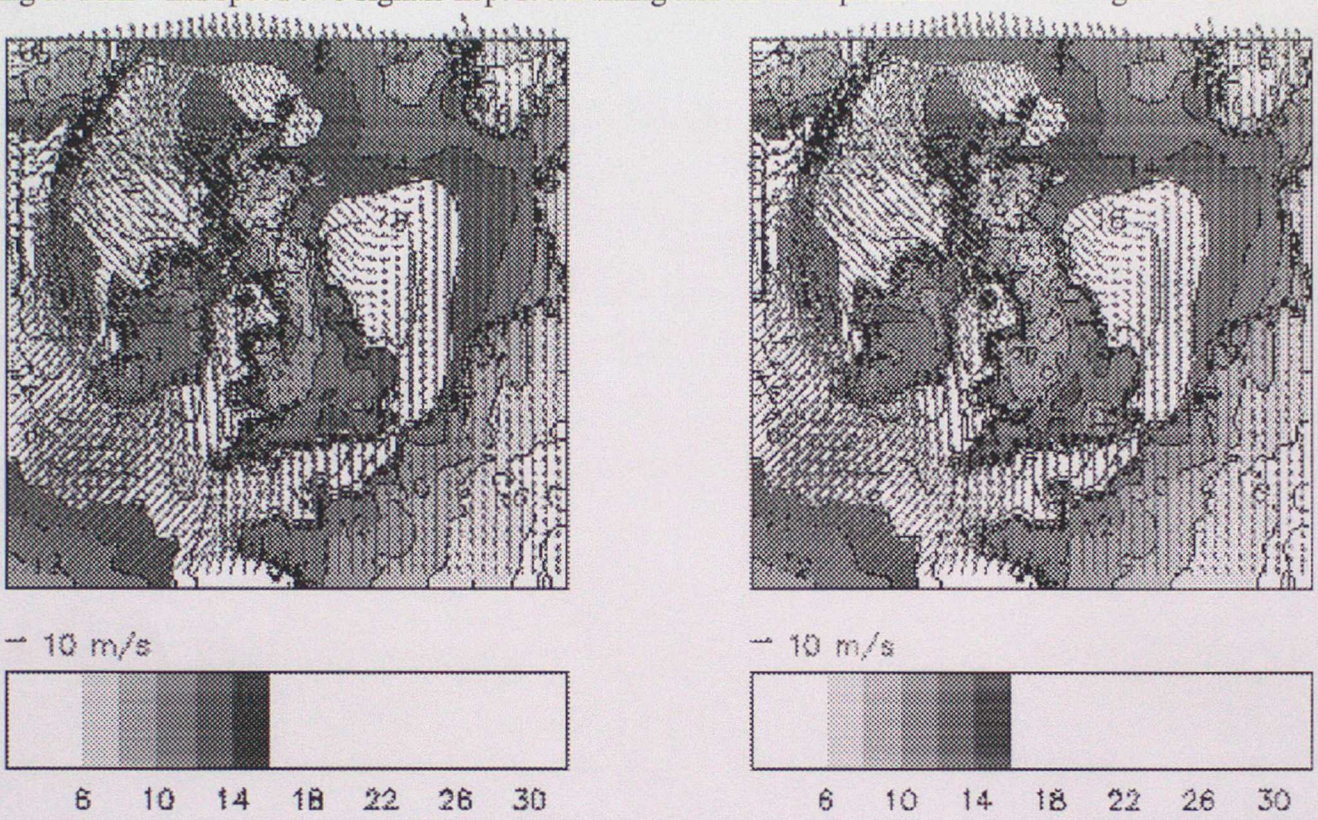

Figure 27: 10m wind speed (ms^{-1}) for T+15 from DT 00Z 17 January 1995. Control is left chart.

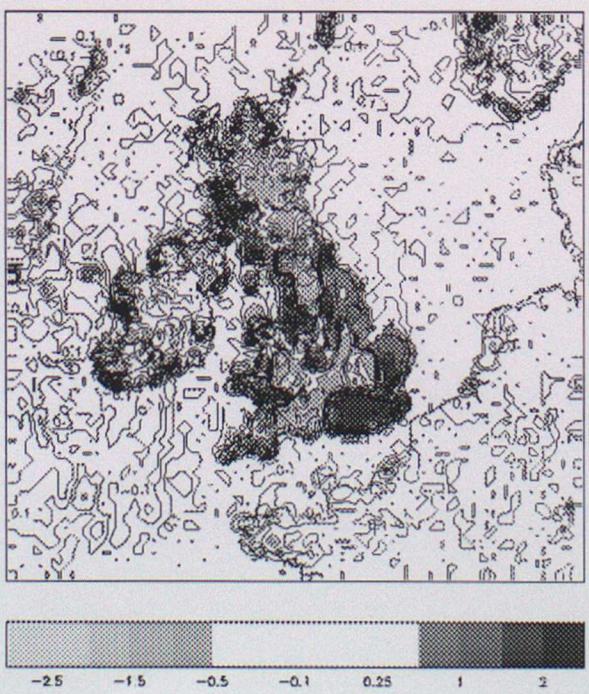

Figure 28: Difference of figure 27, control - ITE. Dark shading shows areas where winds are significantly lighter in the ITE run and light shading shows where winds are significantly stronger in ITE run.

Winds tend to be slightly lighter over the London area and slightly stronger over Scotland. These two signals were present in a number of the experiments although the magnitude of the difference tended to be less in the other experiments. This difference in the 10m wind field can be directly attributed to changes in the roughness length between control and the ITE experiment. The roughness values over London have been increased and the values over Scotland decreased in the ITE datasets compared to control.[1]

Other fields were assessed but no clear signals emerged. It may be expected that a signal would be present in the screen temperature field, particularly in those regions which have seen the introduction of an urban class due to the different thermal properties of urban compared to vegetation. However, even though some significant differences were noted, they were contrary to one another. For example, in the DT 170195 case London was about 0.5°C cooler in the ITE run compared to control throughout much of the forecast but in the DT 111095 case temperatures in the London area were about 0.5°C warmer. Other cases showed no clear signal. Generally, although differences in screen temperature were extensive, they were too small to be considered significant.

Other fields such as total cloud, 1.5m relative humidity, precipitation type, rate and accumulation also all showed no significant impact. There was no impact in the formation, extent or clearance of radiation fog although again there was slight variation in the fog fraction values in those grid boxes where the fog fraction value was less than 1.

The 4 assimilation cycle experiment also showed no significant impact in any field.

Summary of Subjective Assessment

The only significant signals were in the 10m wind field with winds over Scotland slightly stronger and winds in the London area slightly lighter. In other fields although differences were extensive they were largely insignificant.

[1]In an attempt to compensate defficiencies in the WHS data the roughness length values for grass and soil were artificially increased in the control run (as used in the operational model). Furthermore, the vegetation roughness length, as described here, is combined with orographic roughness components within the parametrization to form an effective roughness length. The orographic roughness components were artificially increased over the London area in the control run but returned to their correct values in the ITE run.

Since Scotland is mainly rough grazing (figure 6) the reduction of roughness values for grass and soil is having the greatest effect. The introduction of an urban class increases the vegetation roughness length in the London area making the correction to the orographic roughness components unnecessary. Evenso, the ITE data has higher roughness values in the London area compared to control.

7. CONCLUSIONS

The ITE NATLAC land classification data has successfully been processed and implemented in the Unified Model. The data has introduced the important surface land classes of forestry and urban for the first time in the Mesoscale Model. The data certainly provides a much better depiction of the true surface type over England, Scotland and Wales than was previously the case.

The most significant impact of the data has been in the improved verification of 10m wind. Wind speeds are generally reduced particularly in the London area but winds over Scotland tend to be slightly stronger. Other surface fields showed no significant impact.

It is perhaps surprising that there was not a bigger impact in the screen temperature field, especially over the urban areas. A mixed signal was present in the London area so no conclusions can be drawn. It may be that the resolution of the current Mesoscale Model is still too coarse for London to be resolved or that single cases studies do not provide enough spin-up time for any consistent signal to develop.

The work has shown that it is possible to import data into the system that has been developed using the WHS dataset and is an important prototype for future datasets. The ultimate aim must be to have accurate knowledge of the global surface type/vegetation at a resolution in the order of 1km. It is known that such a dataset is being developed by the US Geological Survey. Datasets of this accuracy will be required for the high resolution models of the future and to accurate model sub-grid scale heterogeneity. Eventually, real-time vegetation index data may also be used.

8. REFERENCES

Ball DF, Radford, GL and Williams WM (1983) 'A Land Characteristic Data Bank for Great Britain'. Bangor Occasional Paper No. 13, Institute of Terrestrial Ecology, Bangor Research Station, Penrhos Road, Bangor, Gwynedd, LL57 2LQ.

Wilson MF and Henderson-Sellers A (1985) 'A Global Archive of Land Cover and Soils Data for use in General Circulation Climate Models'. Journal of Climatology Vol 5 pp119-143.